INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION

INTERNATIONAL CENTRE FOR SCIENCE AND HIGH TECHNOLOGY
c/o INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS 34100 TRIESTE (ITALY) VIA GRIGNANO, 9 (ADRIATICO PALACE) P.O. BOX 586 TELEPHONE 040-224572 TELEFAX 040-224575 TELEX 460449 APH I

H4.SMR/537-20

SECOND COLLEGE ON THEORETICAL AND EXPERIMENTAL RADIOPROPAGATION PHYSICS
(7 January - 1 February 1991)

Co-sponsered by ICTP, URSI, ICSU and with the participation of ICS

PROBABILITY AND RANDOM VARIABLES

S. C. Dutta Roy
Indian Institute of Technology, Delhi
New Delhi, India

PROBABILITY AND RANDOM ~~PROCESSES~~ VARIABLES.

S C DUTTA ROY
Electrical Engineering Department
Indian Institute of Technology
New Delhi 110 016
India

Text of a series lectures to be delivered at the URSI/ICTP sponsored Second College on Radio Propagation Physics, Trieste, Jan 7 - Feb 1, 1991

INTRODUCTION

A random signal is one which has erratic and apparently unpredictable variations. The received signal in a radio communication system is random due to various causes e.g. random interference, receiver noise etc. We cannot predict the exact value of the signal at specified instant. Nevertheless, it can be described in terms of its statistical properties such as average power or ~~the~~ its spectral distribution. The mathematics necessary for this purpose ~~is~~ belongs to the domain of probability theory.

NOTION OF PROBABILITY

Tossing a coin results in "head" or "tail". If this experiment is repeated, the outcome of each is unpredictable. Such an experiment is referred to as a random experiment, which in general, satisfies three features viz repeatability under identical conditions; unpredictable outcome; outcomes averaged over a large number of trials showing some statistical regularity or pattern.

Let event A denote one of the possible outcomes of a random experiment. In n trials of the experiment, let A occur n_A times. Then n_A/n is called the relative frequency of the event A and clearly

$$0 \le n_A/n \le 1 \qquad (1)$$

when $n_A/n = 0$ if A is an impossible event, when $n_A/n = 1$ if A is the only possible outcome.

If n_A/n converges to the same limit as n becomes larger and larger, we say that the experiment exhibits statistical regularity and we define the probability of event A as

$$P(A) = \lim_{n\to\infty} \frac{n_A}{n} \qquad (2)$$

AXIOMS OF PROBABILITY

Let, to each possible outcome of a random experiment, a point called sample point (denoted by S_k) be associated. The totality of sample points is called the sample space, which we denote by S. An event corresponds to either a single sample point or a set of sample points. In particular, a single sample S is called the <u>sure event</u>, the null set ϕ is called the <u>null</u> or <u>impossible event</u>, and a single sample point is called an <u>elementary event</u>.

An example in point is the random experiment of throwing a die. There are six possible outcomes i.e.

$$S = \{1, 2, 3, 4, 5, 6\}$$

S is a one dimensional sample space consisting of 6 sample points as shown in Fig. 1

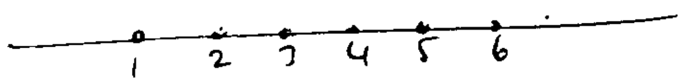

Fig. 1

"A six" ~~shows~~ is an elementary event, while "an even number of dots" ~~shows~~ corresponds to the subset $\{2, 4, 6\}$ of the set S.

We now formally define probability. A probability-system consists of the triple:

1. A sample space S of elementary outcomes
2. A class E of events that are subsets of S
3. A probability measure $P(\cdot)$ assigned to each event A in the class E and having the following properties:

(i) $P(S) = 1$ (3)

(ii) $0 \leq P(A) \leq 1$ (4)

(iii) If $A + B$ is the union of two mutually exclusive events in E, then

$$P(A+B) = P(A) + P(B) \quad (5)$$

Properties (i), (ii) and (iii) are known as the axioms of probability, each of which has a relative frequency interpretation. Axiom (ii) corresponds to (1), while axiom (i) corresponds to the limiting case of (1) when event A occurs in all the n trials. To interpret axiom (iii), let event A occur n_A times in n trials, while event B occurs n_B times. Then the union event A or B occurs in $n_{A+B} = n_A + n_B$ trials; thus

$$\frac{n_{A+B}}{n} = \frac{n_A}{n} + \frac{n_B}{n}$$

which has a mathematical form similar to (5).

ELEMENTARY PROPERTIES OF PROBABILITY

<u>Property 1</u> $P(\bar{A}) = 1 - P(A)$ (6)

where $\bar{A}$ = complement of event A.

<u>Property 2</u> If M mutually exclusive ~~...~~ events $A_1, A_2 \cdots A_M$ have the exhaustive property

$$A_1 + A_2 + \cdots + A_M = S \quad (7)$$

then $$P(A_1) + P(A_2) + \cdots + P(A_M) = 1 \quad (8)$$

when the M events are equally likely, then $P(A_i) = 1/M$, $i = 1, 2, \cdots M$.

<u>Property 3</u> When events A and B are not mutually exclusive, then the probability of the union event "A or B" equals

$$P(A+B) = P(A) + P(B) - P(AB) \quad (9)$$

where P_{AB} is called the probability of the joint event "A and B"

P_{AB}, the joint probability has the following interpretation

$$P(AB) = \lim_{n\to\infty} \frac{n_{AB}}{n} \quad (10)$$

where n_{AB} = no. of times the event A and B occur simultaneously in n trials of the experiment.

CONDITIONAL PROBABILITY

The probability of event B, given that event A has occurred is called the conditional probability of B given A, and is denoted by $P(B/A)$. Assuming that $P(A) \neq 0$, $P(B|A)$ is ~~defined~~ given by

$$P(B/A) = \frac{P(BA)}{P(A)} \qquad (11)$$

To ~~justify~~ prove (11), suppose in n trials, ~~n_{AB} times~~ A and B occur jointly n_{AB} times, and n_A times A occurs. Then $n_A \geqslant n_{AB}$ i.e.

$$\frac{n_{AB}}{n_A} \leq 1$$

The LHS is the relative frequency of B given that A has occurred. For large n, $n_{AB}/n_A = P(B|A)$ i.e.

$$P(B|A) = \lim_{n\to\infty} \frac{n_{AB}}{n_A} = \lim_{n\to\infty} \frac{n_{AB}/n}{n_A/n}$$

$$= P(AB)/P(A)$$

Thus

$$P(AB) = P(B|A)\,P(A) = P(A/B)\,P(B) \qquad (12)$$

The relation

$$\frac{P(B/A)}{P(B)} = \frac{P(A/B)}{P(A)} \qquad (13)$$

is one form of <u>Bayes' rule</u>.

If ~~If~~ events A and B are not in any way linked with each other, then $P(B|A) = P(B)$, $P(A/B) = P(A)$ and (12) becomes

$$P(AB) = P(A)\,P(B) \qquad (14)$$

RANDOM VARIABLES

The outcome of an experiment is customarily thought of as a variable whose value is determined by the outcome. A function whose domain is a sample space and whose range is some set of real numbers is called a random variable (RV) of the experiment. Thus if s is a point in the sample space S, then the RV associated with s is a rule or relationship $X(s)$ that assigns a real number X to every sample point s. If S contains a countable number of sample points e.g. in throw of a die, then $X(s)$ is a discrete RV having a countable number of distinct values. If, however, X can take any value in a finite interval, $X(s)$ is a continuous RV, an example being the amplitude of a noise voltage.

Consider a RV X and the probability of the event $X \leq x$. We denote this by

$$P(X \leq x) = F_X(x) \qquad (15)$$

where $F_X(x)$ is called the cumulative distribution function or simply the distribution function of the RV X. It has the following properties

1. $0 \leq F_X(x) \leq 1$ (16)
2. $F_X(x_1) \leq F_X(x_2)$ if $x_1 < x_2$ (17)

The first derivative of the distribution function $F_X(x)$ is called the probability density function $f_X(x)$:

$$f_X(x) = \frac{d}{dx} F_X(x) \qquad (18)$$

It follows that the probability of the event $x_1 < X \leq x_2$ is

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1)$$
$$= F_X(x_2) - F_X(x_1)$$
$$= \int_{x_1}^{x_2} f_X(x)\,dx \qquad (19)$$

Since $F_X(\infty) = 1$ and $F_X(-\infty) = 0$, it follows from (19) that

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 \qquad (20)$$

Thus,

a probability density function must always be nonnegative and the total area under its curve is unity.

SEVERAL RANDOM VARIABLES

Consider two RV's X and Y. The joint distribution function $F_{X,Y}(x,y)$ is defined as

$$F_{X,Y}(x,y) = P(X \leq x,\ Y \leq y) \qquad (21)$$

If $F_{X,Y}(x,y)$ is continuous everywhere, and the partial derivative

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x\,\partial y} \qquad (22)$$

exists and is continuous everywhere, then $f_{X,Y}(x,y)$ is called the joint probability density function of the RV's X and Y.

Now, by definition, $F_{X,Y}$ is a monotone nondecreasing function of both x and y. Thus $f_{X,Y} \geq 0$, and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1 \qquad (23)$$

Also, note that

$$F_X(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{X,Y}(x,y)\, dx\, dy \qquad (24)$$

Differentiating both sides w.r.t. x, we get

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy \qquad (25)$$

Thus the probability density function $f_X(x)$ may be obtained from the joint probability density function $f_{X,Y}(x,y)$ by simply integrating it over all possible values of the undesired RV, Y. Similarly, for $f_Y(y)$. $f_X(x)$ and $f_Y(y)$ are called marginal densities.

Let X and Y be two continuous RV's with joint probability density function $f_{X,Y}(x,y)$. The conditional probability density function of Y given that $X = x$ is defined by

$$f_Y(y / X = x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \qquad (26)$$

This function satisfies

$$f_Y(y / X = x) \geq 0 \qquad (27)$$

and

$$\int_{-\infty}^{\infty} f_Y(y / X = x)\, dy = 1 \qquad (28)$$

If X and Y are statistically independent, then

$$f_Y(y \mid X = x) = f_Y(y)$$

and

$$f_{X,Y}(x, y) = f_X(x)\, f_Y(y)$$

STATISTICAL AVERAGES

The mean or expected value of a RV X is defined by

$$\bar{x} = m_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx \qquad (29)$$

where E is the expectation operator. Similarly,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx \qquad (30)$$

For the special case $g(X) = X^n$, we obtain the n-th moment of the probability distribution function of the RV X i.e.

$$\overline{x^n} = E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\,dx \qquad (31)$$

$n = 1$ gives the mean value $m_X = \bar{x}$ while
$n = 2$ gives the mean square value $\overline{x^2}$.

We may also define central moments. The n-th central moment is

$$E\left[(X - m_X)^n\right] = \int_{-\infty}^{\infty} (x - m_X)^n f_X(x)\,dx \qquad (32)$$

For $n = 1$, this central moment is of course zero, whereas for $n = 2$, the second central moment is referred to as the variance of the RV:

$$\mathrm{Var}[X] = E\left[(x - m_X)^2\right] = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x)\,dx \qquad (33)$$

The variance is usually denoted by σ_X^2, σ_X being called the standard deviation of the RV X. σ_X is a measure of the dispersion of the variable and is related to $E[X^2] = \overline{x^2}$ and $E[X] = \bar{x} = m_X$ as follows.

$$\begin{aligned}\sigma_X^2 &= E[X^2 - 2m_X X + m_X^2] \\ &= E[X^2] - 2m_X E[X] - m_X^2 \\ &= E[X^2] - m_X^2 = \overline{x^2} - \bar{x}^2 \qquad (34)\end{aligned}$$

Thus if $\bar{x} = 0$ (zero mean RV), σ_X^2 and $\overline{x^2}$ are equal.

JOINT MOMENTS

Consider a pair of RV's X and Y. Joint moments of X and Y are

$$E[X^j Y^k] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^j y^k f_{X,Y}(x,y)\,dx\,dy \qquad (35)$$

when j and k are positive integers. Of particular importance is the case $j = k = 1$; this is called correlation, $E[XY]$. The correlation of $X - E[X]$ and $Y - E[Y]$ is called the covariance:

$$\begin{aligned}\mathrm{cov}[XY] &= E\left[(X - E[X])(Y - E[Y])\right] \\ &= E[XY] - m_X m_Y \qquad (36)\end{aligned}$$

Let σ_X^2 and σ_Y^2 be the variances of X and Y, respectively.

~~12~~ 11

Then the correlation coefficient ρ of X and Y is defined by

$$\rho = \operatorname{cov}[XY] / (\sigma_x \sigma_y) \qquad (37)$$

Two RV's X and Y are uncorrelated iff $\operatorname{cov}[XY] = 0$ and orthogonal iff $E[XY] = 0$. Note, from (36) that if X or Y or both are zero mean RV's and if they are also orthogonal, then they are also uncorrelated. Also note that if X and Y are statistically independent, then they are uncorrelated; however, the converse of this statement is not true.

SOURCE

Simon Haykin, Digital Communications, John Wiley, 1988.